Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physica D 197 (2004) 286–302
PHYSICA D
www.elsevier.com/locate/physd

Pattern formation and nuclear divisions are uncoupled in *Drosophila* segmentation: comparison of spatially discrete and continuous models

Vitaly V. Gursky[a], Johannes Jaeger[b], Konstantin N. Kozlov[c], John Reinitz[b], Alexander M. Samsonov[a]*

[a] *Theoretical Department, The Ioffe Physico-Technical Institute of the Russian Academy of Sciences, 26 Politekhnicheskaya st., St. Petersburg 194021, Russia*
[b] *Department of Applied Mathematics and Statistics, Center for Developmental Genetics, Stony Brook University, Stony Brook, NY 11794-3600, USA*
[c] *Department of Computational Biology, Center for Advanced Studies, St. Petersburg State Polytechnic University, 29 Politekhnicheskaya st., St. Petersburg 195251, Russia*

Received 27 January 2004; received in revised form 18 June 2004; accepted 17 July 2004

Communicated by Y. Nishiura

Abstract

We study the relationship between pattern formation in *Drosophila* segment determination and nuclear structure by replacing the nuclei by a homogeneous continuum. Because this replacement cannot be performed experimentally, mathematical simulation is applied by transforming a previously published model of the segmentation system formulated in terms of explicit nuclear structure into partial differential equations. This transformation has changed the mathematical type of model equations and is therefore interpreted as a structural perturbation of the model. Parameter values are found for three continuum models by means of a new optimal steepest descent algorithm. Each of these models contains a different mathematical representation of nuclear divisions (mitoses). We obtained correct pattern dynamics from all of them, as well as from the model with explicit nuclear structure. This leads us to conclude that *nuclear divisions* are *not coupled to pattern formation* and serve only to populate the blastoderm with nuclei. We also investigate whether the calculated patterns in the developmental period modelled resemble

* Corresponding author. Tel.: +7 812 2479352; fax: +7 812 2471017.
E-mail addresses: samsonov@math.ioffe.ru (Alexander M. Samsonov); gursky@math.ioffe.ru (Vitaly V. Gursky); jjaeger@ic.sunysb.edu (Johannes Jaeger); kozlov@spbcas.ru (Konstantin N. Kozlov); reinitz@ams.sunysb.edu (John Reinitz).

0167-2780/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.physd.2004.07.003

their attractors, and find that they fail to do so. The implications of our results for models of biological pattern formation based on partial differential equations are discussed.
© 2004 Elsevier B.V. All rights reserved.

PACS: 87.10.+e; 87.16.Yc; 87.18.-h; 87.18.Hf

Keywords: Pattern formation; Regulatory gene networks; Nonlinear reaction-diffusion; Nuclear division; Drosophila segmentation

1. Introduction

Animal development is a process that unfolds over time in accordance with certain rules, and hence it is natural to use dynamical equations to model it. A fundamental difficulty facing such an effort is the extreme complexity of biological systems, which includes not only a large number of biological regulators, but also spatiotemporal structure. In general, biological systems contain spatial structure at all scales, from organs down to the molecular level. This fact complicates efforts to construct theoretical models of spatially controlled processes such as development. It is in strong contrast with classical reaction-diffusion theory, such as that used for the Belousov–Zhabotinsky reaction, in which there is a large gap in scale between the molecular mean free path and the scale of structures formed by the balance between reaction and diffusion. It is thus unclear if nonlinear reaction-diffusion equations on a continuum can be used in a biological system, because compartmentalization within and between cells by membranes and other structures invalidates the physical assumptions on which reaction-diffusion theory is based. It is common knowledge in developmental biology that pattern formation implies a necessity of cell (or nuclear) division. More generally, there is no guarantee that animal development can be understood without taking into account the physical structure of a biological system at all scales.

Here, we consider this issue in the context of a specific biological system, segment determination in the early embryo of the fruit fly *Drosophila melanogaster*. At the earliest stages of development, this system is not composed of tissue at all, but rather relatively isolated cell nuclei which divide according to a specific schedule ultimately producing epithelial tissue. We ask whether this tissue formation process is coupled to pattern formation, or independent of it. This is a basic biological question which is refractory to experimental analysis because the nuclear structure of the embryo cannot be disrupted while maintaining the living state. For that reason we analyze the system using mathematical and numerical methods, taking as our starting point dynamical equations directly based on experimental observations.

We perform our analysis by altering these dynamical equations in ways that represent the abolition of the embryo's cellular structure and its replacement by a continuum. Because this alteration leads to a change in the type of equations, we can say that we structurally perturb the model and investigate if it is stable against such perturbation. This procedure is designed to answer the biological question, but is informative at a theoretical level as well. There is a strong contrast between Turing's idea [1] that the determination of biological pattern can be described by reaction-diffusion equations operating in a continuum and the central role of cells and genes in this process as elucidated by modern cell and molecular biology. This raises provocative questions about the validity of Turing's approach in the context of current knowledge. Another question concerns the mathematical representation of cell fate. The stability of determined cells against perturbations such as transplantation indicates that their cellular states are in dynamically stable basins of attraction. In the absence of directly observable state variables, many theories were developed in which the experimentally evident basins of attraction were mathematically represented as stable attractors [2,3]. We will examine the validity of this assumption in a well characterized dynamical model based on the direct monitoring of biological state variables.

This paper studies a mathematical reformulation of a previously proposed model [4,5], formulated in terms of discrete nuclei and incorporating protein *synthesis*, *diffusion* and *decay*. The mathematical representation chosen for that model is a hybrid system in which a mitosis is represented by shutting off protein synthesis for the duration

of the mitosis, followed by a discontinuous doubling of the number of nuclei. During interphase, when protein synthesis is taking place, the system is described by a set of differential-difference equations for the concentrations of segmentation gene proteins. Protein synthesis is controlled by regulation in the zygotic gene network and by a maternal morphogen gradient. Diffusion is modelled by a *discrete* term representing the interaction of a nucleus with its nearest neighbors. In this study we adopt the continuum hypothesis and investigate whether the resulting equations are capable of forming correct gene expression patterns. In what follows, we call the previous model [5] the hybrid model, while the partial differential equations (PDEs) formulated here are called the continuous model. This is to emphasize that we actually deal with two representations of the same biological system.

Adopting the continuum model turns the differential-difference equations into nonlinear reaction-diffusion PDEs. In particular, this implies that the old representation of mitosis is not applicable to the new equations. By considering different ways of representing mitosis we can explore the role of mitosis in the pattern formation process in a manner that is infeasible by direct experiment.

The paper is organized as follows. First, we describe the biological system under consideration with the model equations and provide information on how the parameters in the equations have been found so as to give an optimal fit of the solution to independently obtained expression data. Then, we show the optimal solutions to the model and discuss possible biological issues from these numerical results. Finally, we investigate the solution behavior at large times in the continuous and hybrid versions of the model and discuss the main results. The appendix contains a description of the optimization algorithm and most of the mathematical details of its application to our model.

2. Model

2.1. Biological system

We consider the genetic system that controls segment determination in the fruit fly *Drosophila*. Segments are repeating units that form the body of the fly. The development of segments is controlled by the segmentation genes. Segment determination takes place early in embryogenesis. Immediately following the deposition of a *Drosophila* egg, a rapid series of 13 almost synchronous nuclear divisions take place, without the formation of cells. Between the sixth and ninth nuclear divisions, the nuclei migrate to the outside (cortex) of the egg and form an approximately ellipsoidal shell of cells known as the syncytial blastoderm. The period between the end of nuclear division $C-1$ and nuclear division C is referred to as cleavage cycle C. Cleavage cycle 14A is not terminated by a nuclear division; during this cycle cell membranes invaginate between the nuclei and seal them off into cells. At about the time cellularization is complete, a complex set of folding motions called gastrulation begins. It is during cleavage cycle 14A that segment determination takes place (see Fig. 1).

There are about 40 segmentation genes in *Drosophila* [6–10], divided into four classes: maternal coordinate genes, gap genes, pair-rule genes, and segment polarity genes. This classification was originally based on the phenotypes of mutants. The segmentation genes are expressed in patterns characteristic of the class that they belong to. While there are many maternal coordinate genes, their actual input to the segmentation system takes place via the products of three genes, *bicoid* (*bcd*), *caudal* (*cad*), and *hunchback* (*hb*), which are expressed as three monotonic concentration gradients of protein at the time when the zygotic segmentation genes are first activated. Of these three genes, *bcd* is entirely maternal while both *cad* and *hb* are expressed from both the maternal and zygotic genomes. Five gap genes are expressed in one or two broad domains about 10–20 nuclei wide which gradually intensify and sharpen during the blastoderm period. The seven pair-rule genes are initially expressed almost uniformly but late in the blastoderm period resolve to seven distinct stripes 3–4 nuclei wide. Examples of maternal, gap, and pair-rule expression patterns are shown in Fig. 1. The segment polarity genes are not expressed until gastrulation, where they appear as 14–17 single cell wide stripes. Two segment polarity genes, *engrailed* and *wingless*, which are expressed in adjacent rows of cells, are the final output of the segmentation system. Their expression is stable through the life

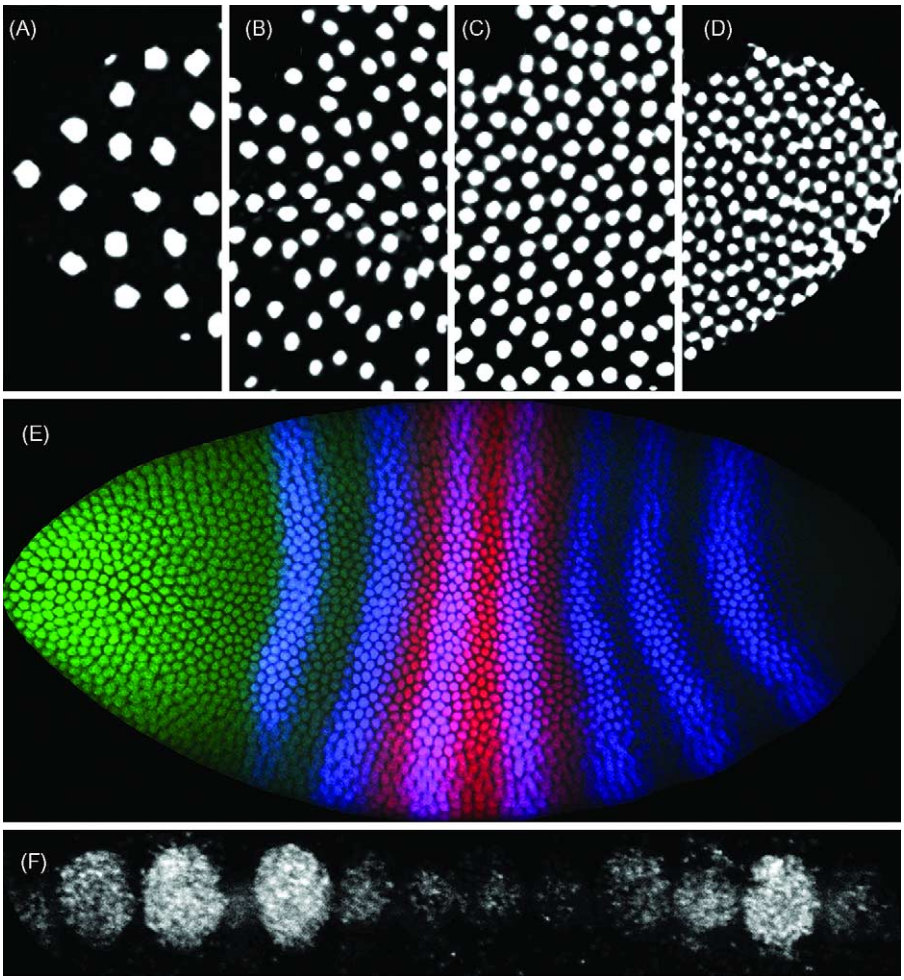

Fig. 1. A–D show the first through fourth quarters along the A–P axis of embryos in cycles 10–13 respectively. These embryos are stained with mouse anti-histone antibody visualized with biotinylated anti-mouse serum and strepavidin conjugated with alexafluor700. The nuclei appear as dots; the apparent size change is a consequence of a shape change. Panel E shows a cycle 14A embryo stained for representatives of the three classes of segmentation genes expressed in the blastoderm. Protein expressed from *bcd*, a maternal coordinate gene, is shown in green. Protein expressed from the gap gene *Krüppel* (*Kr*) is shown in red, and protein expressed from the pair-rule gene *eve* is shown in blue. In panels A–E the outer surface of the embryo is shown. Panel F shows a high magnification cross section of an embryo stained for *eve*. Note the spotty "punctate" staining: this is probably a consequence of Eve protein binding preferentially to certain spots on the chromosome. The embryos are fluorescently labeled with polyclonal antibodies. The serums were raised and the embryos were fixed and stained as described [36]. Anterior is to the left and dorsal is up. Each embryo is about 0.5 mm long.

of the fly and forms a prepattern with segmental periodicity which underlies segment formation at later stages of embryogenesis [11–13].

The blastoderm has certain features which make it uniquely favorable for theoretical studies of development. First, the expression of segmentation genes is to a very good level of approximation a function only of distance along the anterior–posterior (A–P) axis (the long axis of the embryo ellipsoid). This allows us to use models with only one spatial coordinate along the A–P axis. Because the blastoderm is a syncytium, cell-cell interactions can be treated in terms of the diffusion of protein products of genes. Most importantly, the state of the system is

given by the concentrations of protein products of segmentation genes. Prior to gastrulation, nuclear divisions and morphology are under maternal control, as shown by the absence of zygotic mutants which affect these processes [14,15]. However, maternal inputs are not sufficient to set the expression patterns of segmentation genes. These genes are coupled to each other prior to gastrulation, since mutations in one segmentation gene typically affect the expression of others [16], but do not affect embryo morphology until after gastrulation. Hence a model should include regulatory interactions among the segmentation genes, but need not account for the coupling between the segmentation genes and embryo morphology.

2.2. The dynamical equations

Here we will consider *even-skipped* (*eve*) stripe formation in a PDE (continuous) context, using an earlier analysis of a hybrid system [5] as a starting point. The original motivation of the hybrid model was to capture the discrete cellular growth and structure of an embryo [4]. This mathematical formulation has proved biologically useful [5,17,18], but it amounts to a strong coarse-graining of biological structure, in which nuclei with a complex internal structure and their surrounding cytoplasm are approximated by points of no spatial extent separated by a finite distance. A continuum model is also a strong coarse-graining, but of a complementary kind in which the fundamental cellular structure (see Fig. 1) is neglected. The presence of biological structure at all scales in living tissue means that the validity of equations describing a developmental process should be cross-validated over different spatial approximations to give biological insight into the role of spatial structure in pattern formation. That is one motivation for considering the PDEs. An additional one is the possibility of mathematical characterization of the solutions in a biologically meaningful way [19].

We briefly review essential points discussed in detail elsewhere [4,17,19]. We represent each gene phenomenologically in terms of how its rate of protein synthesis depends on the local concentration of various regulators, without attempting to represent the explicit state of active chromatin at the molecular level. The state variables (the unknowns) are the concentrations of segmentation gene products (proteins). The change in time of concentrations of these products is governed by three processes: regulated protein synthesis, diffusion, and first order decay. We consider an interval of length L on the A–P axis and introduce the continuous variable $x \in L$. Denoting the concentration of the ath gene product in a nucleus at position x at time t by $v^a(x,t)$, we write a set of coupled nonlinear reaction-diffusion equations for N zygotic genes as:

$$\frac{\partial v^a(x,t)}{\partial t} = R^a(t) g\left(\sum_{b=1}^{N} T^{ab} v^b(x,t) + m^a v^{\mathrm{Bcd}}(x) + h^a\right) - \lambda^a v^a(x,t) + D^a \frac{\partial^2 v^a(x,t)}{\partial x^2}, \tag{1}$$

where $a = 1, \ldots, N$.

The first term on the right hand side of (1) describes the regulated rate of synthesis of protein from the ath gene. The function $g(\cdot)$ is a monotonic sigmoid ranging from zero to one, and its formal description is not unique. Here we use $g(u^a) = (1/2)[(u^a/\sqrt{(u^a)^2+1}) + 1]$ for all a, where $u^a = \sum_{b=1}^{N} T^{ab} v^b + m^a v^{\mathrm{Bcd}} + h^a$. The regulation of gene a by gene b is characterized by the element T^{ab} of matrix $\mathbf{T}$. T^{ab} is a nonlinear sensitivity coefficient in the sense that if $g(u^a) = 0.5$, a small change δv^b in regulator v^b results in a change of $T^{ab}\delta v^b$ in the synthesis rate. For other values of u^a the size of the change will be less, and may approach zero. The term $m^a v^{\mathrm{Bcd}}(x)$ describes the regulatory effect of the maternal Bcd protein, which varies in space but is constant in time. Other maternal transcription factors are constant in both space and time, and their effect on gene a is summarized by h^a. The maximum rate of synthesis for protein a is given by the function $R^a(t)$, which depends on time because of the effects of mitosis. The form of this dependence is discussed below. The second term on the right hand side of Eq. (1) describes the degradation of ath protein, which is modelled as first order decay with rate λ^a. Diffusion of protein in the embryo is described by Fick's law and taken into account with the last term on the right hand side of (1). The distribution of nuclei is assumed to be homogeneous (see Fig. 1), and, hence, the diffusion rate is spatially constant.

In what follows we apply Eqs. (1) to a network of $N = 5$ genes, including the gap genes (*Kr*, *hb*, *gt*, and *kni*) and the pair-rule gene *eve*. Our model covers the central part of the embryo including *eve* stripes 2–5 from cleavage cycle 11 to cycle 14A. We introduce a dimensionless form for (1), and rescale x by L, so that $x \in [0, 1]$, and time by a scale of 1 min for the whole time period equal to approximately 88 min [20]. Experimental data are available for this network, and concentrations for all genes are scaled against a maximal expression level of 12 as described [17,5].

The initial and boundary conditions are as follows. Because zygotic gene expression begins at cycle 11, all purely zygotic genes have zero concentrations for initial conditions. Gene *hb* is expressed from both the maternal and zygotic genomes, and its maternal expression pattern, a sigmoid gradient falling from anterior to posterior, is used as a nonzero initial condition for that gene. The boundary conditions are formalized by zero flux condition at the ends of the spatial domain, which does not cover the whole embryo here. The actual flux will be nonzero and dependent on gene expression in a complicated manner, so we choose zero flux because other numerically feasible alternatives, such as periodic boundary conditions, are not biological.

The Eqs. (1) are a continuous reformulation of the earlier model equations [4,5]. Those equations were a hybrid system in which interphase was described by a *differential-difference* equation version of (1), mitosis by (1) with $R^a = 0$, and the cleavage of nuclei at the end of mitosis by a discrete transition, in which the number of nuclei, and hence, the number of equations doubled. In that hybrid system, the protein synthesis rate was a property of a single nucleus. Similarly, the diffusion rate in the hybrid system was considered to scale with the inverse squared distance between two nuclei, and thus quadrupled in each nuclear division as the nuclear distance was halved (in one dimension). In (1), D does not depend on nuclear cleavage, but we are left with the question of the proper formulation for $R^a(t)$.

Indeed, the choice of $R^a(t)$ is a nontrivial theoretical issue. If the nuclei were smaller in size than any spatial scale in which gene expression changes, and if macromolecular synthesis took place in a region of infinitesimal spatial extent, writing (1) would be a straightforward exercise in taking concentrations, and $R^a(t)$ would double in each successive cleavage cycle. In fact, this approximation does not hold (see Fig. 1). This is both because nuclei are large compared to the scale of spatial variation and because the actual process of protein synthesis takes place in a volume larger than a nucleus, since RNA must be transcribed and processed in the nucleus and transported to the cytoplasm for translation. Newly translated protein returns to the nucleus to bind to chromatin and regulate other genes. A central problem in modelling biological systems is the extent to which various mechanisms must be incorporated into a model of a particular process in order to correctly understand its behavior. Here we are concerned with the spatial part of this coarse-graining problem, which is to ask which (if any) formulations of $R^a(t)$ allow the blastoderm to be well represented by PDEs.

In this paper we consider three possible formulations of $R^a(t)$, where $R^a(t) \geq 0$. These represent different approximations of this dependence, and in order of increasing complexity are as follows:

(A) $R^a(t) = R^a$, a constant.
(B) $R^a(t) = 0$ during mitosis and has a positive value $R^a(t) = R^a$ during interphase.
(C) $R^a = 0$ during mitosis and $R^a(t) = 2^{C-14}R^a$ during interphase, where C is a number of cleavage cycle and R^a is the cycle 14A synthesis rate. This is the same as B, but $R^a(t) \mapsto 2R^a(t)$ in each successive cleavage cycle.

The mitosis schemes A–C reflect different ways of incorporating mitosis into the model. Their ability to reproduce expression patterns (or lack thereof) allows us to draw conclusions about the importance of mitosis to the pattern formation process. At a mathematical level they represent three independent modifications of the hybrid model [5].

3. Fitting procedure

The parameters of the model are found by fitting the numerical solution of (1) to gene expression data. We use the same data as was used in an earlier study of the hybrid model [5] (gap gene and maternal data therein is from

[17]). This data set contains the protein concentrations for segmentation genes at several times during cleavage cycles 13 and 14A, where for each time the data takes the form of a one dimensional row of discrete nuclei running in an A–P direction.

The fitting procedure consists in minimization of the functional

$$F = \sum_{\text{genotypes}} \sum_{a} \sum_{i} \int_{L} (v^a(x, t_i) - v^a_{\text{data}}(x, t_i))^2 \, \mathrm{d}x + P(q), \tag{2}$$

where $v^a(x, t)$ is the numerical solution of the model equations, $v^a_{\text{data}}(x, t)$ is a spatially continuous function obtained by spatial interpolation of data, and $P(q)$ is the penalty function described below. The innermost summation is performed over all time points t_i at which we have data, the next sum is over all genes in the network enumerated by a, and the outermost sum is over two genotypes, wild type and *eve* mutant data. The integration is performed over the spatial domain L. We consider F to be a function of an $N(5 + N)$-dimensional vector of parameters, $q = (R^a, T^{ab}, m^a, h^a, D^a, \lambda^a)$, and we find an optimal parameter set q_{opt} at which F takes its minimal value. Some of parameters q are subjected to inequality constraints for biological reasons. The gap gene data is from three different times t_i in cleavage cycle 14A, while the *eve* data is from two time points in cleavage cycle 13 and two in cycle 14A. The *eve* data in cycle 13 has the form of spatially constant concentration values accounting for the fact that *eve* pattern breaks from an almost homogeneous state to seven stripes only in mid cleavage cycle 14A.

The penalty function $P(q)$ in (2) is defined as proposed in [5]:

$$P = \begin{cases} e^{\Lambda Q} - e, & \text{if } \Lambda Q > 1, \\ 0, & \text{otherwise;} \end{cases}$$

where

$$Q = \sum_{a,b} ((T^{ab} y^b_{\max})^2 + (m^a y^{\text{Bcd}}_{\max})^2 + (h^a)^2),$$

and $y^b_{\max}$, $y^{\text{Bcd}}_{\max}$ are the maximal values of the gene expression data for gene b and Bcd respectively. The penalty function with the parameter Λ is used to control the size of the search space for parameters that contribute to the argument of the function g in the model equations, since the same parameters appear in Q. A very large penalty will arise whenever Q is greater than $1/\Lambda$, and we have chosen $\Lambda = 10^{-7}$ in calculations.

The method of minimization is an optimal steepest descent algorithm described in detail elsewhere [21] (see Appendix A). This algorithm is based on a Lagrangian approach, in which the functional (2) is minimized subject to the constraints that (1) is satisfied and that the parameters lie within their search space. Search space constraints, initially expressed as inequalities, are transformed into additional constraint equations. An expanded functional (Lagrangian) is constructed by adding each constraint equation multiplied by its Lagrange multiplier to (2), and minimized by steepest descent [22]. This method requires a starting point, for which the parameters of the hybrid model were used to obtain the solution to model B. This in turn was used as a starting point to obtain solutions to models A and C.

We believe that the results presented below in fact represent a very close numerical approximation to the global minimum of (2) for models A–C. In order to increase our confidence as to that point, we also performed over 10^5 random perturbations of the optimal parameter values, in which each parameter had a random number added to it of a magnitude of 40% or less of that parameter's value. In no case was a smaller value of (2) found than what was obtained by optimal steepest descent.

4. Optimization results and expression patterns

The expression patterns obtained by the fitting procedure for the three mitosis models A–C described above are shown in Fig. 2. The figure compares gene expression data and model solutions at two time points, one in early and one in late cleavage cycle 14A. The patterns look very similar to those obtained in the hybrid version of the model [5]. The patterns of models A and B have nearly the same rms (root-mean-squared difference between data and the model solution; model A: rms = 1.04, model B: rms = 1.00, hybrid model: rms = 0.927), while the rms for model C is somewhat larger (1.28). The largest deviations between model and data for model C are in early cleavage cycle 14A, particularly for *hb* and *gt*. At the same time, the amplitude of *eve* stripes at late cycle 14A is closer to the data values than in the mitosis models A and B.

We have checked that all of the biological results reported for the hybrid model [5] are also found in models A–C. Specifically, we obtain the same structure of the regulatory matrix **T**: gap genes show cross repression, repression of *eve* and self-activation; the regulation of gap genes by *eve* is close to zero; and diffusivity of *eve* is close to zero (see Appendix A). At a finer level of detail, we find that each of the eight borders of stripes 2 through 5 are under repressive control by a single gap gene, each as previously reported. Finally, we find that the timing of stripe formation is controlled as described [5], both overall and with respect to the earlier appearance of stripe 2.

This result is not a trivial consequence of the model's ability to fit any pattern, since *the hybrid model fails* to make a pattern that fits data if known genes are omitted [17], or if *eve* stripes in the data are displaced by as little as one nucleus [18].

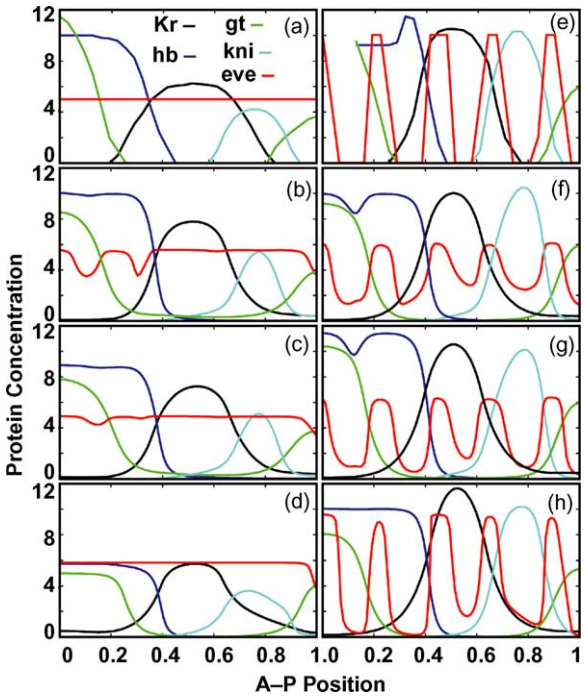

Fig. 2. Segmentation gene expression patterns: comparison between data and models A–C. Protein concentration profiles are shown at early (a–d) and late (e–h) cleavage cycle 14A: (a and e) data, (b and f) model A, (c and g) model B, (d and h) model C. The horizontal axis represents the rescaled spatial domain (covering the middle 32% of the A–P axis of the embryo), the vertical axis represents protein concentrations in conventional units.

These observations indicate that the biological results of the analysis are not sensitive to the way that mitosis is represented in the model. Not only do both the hybrid model and the continuous model with mitosis governed by models B and C give the same results, but model A, in which mitosis is neglected entirely, also does so. Indeed, model C, which has the most elaborated representation of mitosis, gives the poorest fit. This insensitivity to the representation of mitosis indicates that *the early nuclear cleavages play no essential role in the pattern formation process*, but merely serve to populate the blastoderm and generate an epithelium in which pattern formation can take place.

5. Long term dynamics

It is an implicit assumption in many theoretical studies of development and gene regulation that the final patterns observed experimentally are matched by mathematical attractors in corresponding models. In particular, some modelling efforts have represented the gap gene domains as stable attracting states [23]. Here we have simply fit trajectories and not made any assumptions about long time behavior. We can thus use the model to find out if the patterns occurring in the model at the final time when gastrulation starts are asymptotically stable and provide actual attractors of our dynamical system or states which are very close to them. This may shed light on the validity of approximating segmentation gene expression patterns by attractors.

Fig. 3 shows the solutions for models A–C and the hybrid model at late times ($t = 2000$; gastrulation is at $t = 88$). Solutions of models B, C, and the hybrid one settle to a stationary state; the solution of model A oscillates in time in the anterior 25% of the spatial domain and is stationary elsewhere as shown in Fig. 4. If it occurs, stationary behavior is evident at $t = 500$–1000. The differences in long-term patterns among models A–C are a result of the interplay between parameter values and implementations of mitosis, which vary between the models.

Note that the asymptotic patterns do not coincide with the patterns at $t = 88$. This is particularly true of the *eve* stripes, as they are downstream from the gap genes in a regulatory sense and are very sensitive for this reason to changes in the gap gene patterns. Therefore, the patterns in the model at gastrulation are quite far from the genuine attractors.

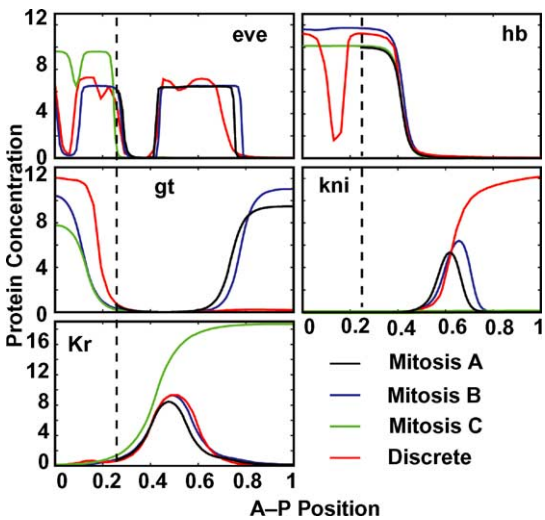

Fig. 3. Asymptotically attracting patterns: model solutions at $t = 2000$ are shown for model A (black), B (blue), C (green), and hybrid model (red) for all genes. Horizontal and vertical axes are as in Fig. 2. The vertical dashed line on the panels separates the spatial domains with two different types of long-term behavior of the model A solution: the solution is stationary in the posterior domain and oscillatory in time in the anterior domain (its graph is not shown in this domain).

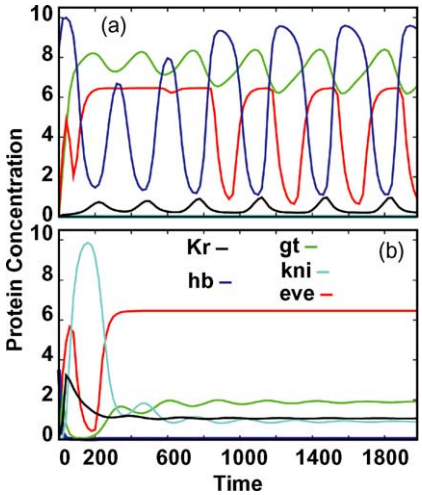

Fig. 4. Long-term dynamics of protein concentration patterns for model A at spatial positions $x = 0.15$ (a) and $x = 0.70$ (b). The horizontal axis represents time in minutes, the vertical axis represents protein concentrations in conventional units. Gastrulation time is at $t = 88$.

6. Discussion

6.1. Biological implications

Gene circuit models are unable to fit data which has been altered from experimental reality in relatively minute ways, such as a shift in expression pattern by as little as a single nucleus [18]. In the cited work, it was established that testing the ability of the model to fit or not fit a given set of data is itself a scientifically informative procedure. Here we use the closely related procedure of investigating which of a group of closely related models can fit the data.

The fact that the biological behavior of a continuous reaction-diffusion model is the same as a hybrid model with nuclear divisions included explicitly indicates that these nuclear divisions have no role in the pattern formation process. This conclusion is amplified by the fact that the PDEs show this property if mitosis is neglected (model A), included as a shutting off of synthesis only (model B), or included as a shutting off of synthesis followed by a doubling of synthetic density (model C). This result could not have been obtained experimentally, since it is impossible to replace the cellular structure of the blastoderm by a continuum in which the segmentation genes act. Nevertheless, it is possible to produce mosaics in which nuclear density is reduced in portions of the embryo, and *eve* stripes nevertheless form [24].

Note that the sets of parameter values in models A–C, and in the hybrid model are quantitatively different from each other, but are qualitatively equivalent (see Appendix A.3). Therefore, they represent the same genetic regulatory system (or the same gene network topology) which is independent of the representation of subcellular structure and the implementation of mitosis in the model.

More generally, we have shown that the biological conclusions drawn from a gene network model [5] are robust against changes in the spatial representation used. This has important implications for modelling biological systems. It contrasts, for example, with a classical fluid representation in which the Navier–Stokes equations are valid for all scales large compared to the mean free path of a molecule. As theoretical studies of experimentally characterized biological systems enter the mainstream, it becomes important to establish the domain of validity of particular modelling frameworks.

Theoretical analysis requires the decomposition of a biological system into relatively weakly interacting modules. There is no guarantee that this is even possible. For example, although specific genes are known to control

development, there is no reason a priori, why these specific genes might not be required to couple properly, e.g., with intercellular variations in the intracellular arrangement of mitochondria in order to function properly. Such a possibility amounts to a 'coupling catastrophe', in which biological processes cannot be understood until their constituent parts will be understood at all levels. Here we have shown that segment determination in the *Drosophila* blastoderm is not coupled to cellular structure, demonstrating that at least one form of the 'coupling catastrophe' is absent in this system.

6.2. Implications for pattern formation models

6.2.1. Turing models

The use of PDEs for modelling development originates in the seminal paper of Turing [1]. Turing considered "masses of tissues which are not growing, but within which certain substances are reacting chemically, and through which they are diffusing" (ibid, p. 38). This level of Turing's approach is well supported by results presented here. The central problem Turing addressed in his paper was whether patterns could arise from a uniform initial state. To answer this question, he considered for parsimony a pair of substances with differing diffusion rates whose reaction term was in a stable state in absence of diffusion. Nevertheless, he was aware that the *biological* problem in pattern formation was different, since as he pointed out "Most of the organism, most of the time, is developing from one pattern into another, rather than from homogeneity into a pattern." (ibid, pp. 71–72).

Other investigators have used Turing's assumptions to model particular biological systems by "Turing Models", describing pattern formation as a diffusion-driven instability from a homogeneous state to a heterogeneous pattern (Turing's instability). Many such models consider the *Drosophila* segmentation system [25–31]. While the work reported here supports PDE models of the *Drosophila* blastoderm, we call attention to the fact that the biologically well supported system of PDEs characterized in this paper does not begin from a homogeneous initial state, and furthermore the initial state is not rendered unstable by diffusion. Pairs of differentially diffusing substances are not evident, and there is no point in the development of the organism where a linear analysis can be employed. These are not merely modelling choices we have made, but also reflect fundamental biological properties of the system.

These points appear to indicate that many "Turing Models" (although not Turing's actual work !) are not applicable to well characterized systems and to the *Drosophila* segment determination system in particular. They may remain important in less characterized systems. For example, it is significant that a "Turing Model" proposed by Meinhardt was able to predict correctly an obligatory coupling of A–P and dorso-ventral patterning genes in limb development [32,33], even though the original model was formulated with no information about biological state variables. It shows that generic pattern formation mechanisms exist, and "Turing Models" may represent very often the simplest examples of such mechanisms, providing a useful form to characterize and enumerate them [26].

6.2.2. Attractors

Our observation that attraction states at long times are quite different from those seen at gastrulation also has theoretical implications. It is clear experimental evidence for basins of attraction in the *Drosophila* segmentation system. For example, heterozygotes for gap mutants have a normal number of *eve* stripes whose spacing is slightly changed and survive [34], but gap mutant homozygotes have a grossly altered number of stripes and die [35]. The heterozygotes are evidently in the same basin of attraction as the wild type, but homozygous mutants are in a different, and fatal, basin. However, the attracting states of these basins are properties of a much larger dynamical system, since thousands of additional genes turn on at the time of gastrulation. Therefore, the attracting states of the blastoderm segmentation system are unobservable, while their basins are not only observable but biologically very important. This fact implies that the key objects for dynamical analysis of segmentation genes are the separatrices between basins of attraction, not the attractors themselves. We do not believe that this fact means that the very useful mathematical properties of asymptotic states cannot be used in the analysis of this system, but rather that any use of them must be in close contact with the biologically observable basin structure.

Concerning quantitative behaviour of parameter set, our preliminary numerical study reveals that the patterns in the models remain reasonably good in a small vicinity of optimal values for parameters (2–5% of the values). The robustness of networks found will be studied more systematically in the future.

Acknowledgements

The support of this work by NIH Grants RR07801 and TW01147, and by CRDF GAP Awards RBO685 and RBO1286 is gratefully acknowledged. We thank Dr. Carlos E.V. Alonso and David Kosman for confocal scans.

Appendix A. Fitting procedure

A.1. Outline of the optimization algorithm

Here we describe the optimal steepest descent algorithm (OSDA) used to fit the model to gene expression data. Assume we have a model formulated as a system of ordinary differential equations with some initial condition:

$$\frac{\mathrm{d}v}{\mathrm{d}t} = f(v, q), \qquad v(0, q) = V, \tag{A.1}$$

where the vector function $v(t, q) = (v_0(t, q), \ldots, v_{K-1}(t, q))^{\mathrm{T}}$ describes the system at time t and $q = (q_0, \ldots, q_{I-1})^{\mathrm{T}}$ is a vector of parameters, where K is the number of state variables, I is the number of parameters, and index T denotes transposition of the vector. We will not discuss the simple case when we have an analytical solution of this system. In general however, many problems can only be formulated in a form of system (A.1) which cannot be solved analytically.

Our aim is to find parameter values providing a good fit of a model solution v to some experimental observations ('data') detected at time moments t_i, $i = 1, \ldots, J$ and denoted as $y(t_i) = (y_0(t_i), \ldots, y_{K-1}(t_i))^{\mathrm{T}}$. In what follows, we use the following abbreviations for the vectors $v(t_i, q)$ and $y(t_i)$: $v_i = v(t_i, q)$, $y_i = y(t_i)$, $i = 1, \ldots, J$.

To characterize the "distance" between solution and data, we introduce the quality functional

$$F(v, q) = \sum_{i=1}^{J} (v_i - y_i)^{\mathrm{T}} (v_i - y_i) + P(q), \tag{A.2}$$

where $P(q)$ is a penalty function. A parameter set q_{opt} is called an optimal parameter set, if it provides a minimal value of the functional F so that

$$F(v, q_{\mathrm{opt}}) = \min_q F(v, q).$$

Here we apply the Lagrange approach for minimization of the functional F. However, there are additional constraints imposed on certain parameters q_i such that we require

$$q_i^{\mathrm{low}} \le q_i \le q_i^{\mathrm{up}}. \tag{A.3}$$

These constraints represent *search space limits* based on a priori knowledge about the process described by the model which need to be incorporated into functional F. This is conveniently achieved by introducing new parameters r_i connected with the original parameters through the trigonometric transformation

$$q_i = \alpha_i + \beta_i \sin(\gamma r_i), \tag{A.4}$$

where γ is an arbitrary constant and α_i and β_i are defined to be

$$\alpha_i = \frac{q_i^{\mathrm{up}} + q_i^{\mathrm{low}}}{2}, \qquad \beta_i = \frac{q_i^{\mathrm{up}} - q_i^{\mathrm{low}}}{2}.$$

It is evident that the q_i in (A.4) obey the inequalities (A.3) for any values of γ and r_i. Note that the trigonometric transformation for parameters under constraints is not unique, and, for instance, one may also apply a hyperbolic tangent, so that

$$q_i = \alpha_i + \beta_i \tanh(\gamma r_i). \tag{A.5}$$

We insert representations (A.4) or (A.5) into system (A.1). In what follows, we will rename the new parameters r_i back to q_i.

In order to incorporate the differential constraint (A.1) (i.e., the model equations), which now includes all constraints (A.3) into the quality functional F, we introduce the expanded quality functional

$$L = F(v, q) + \sum_{i=0}^{J-1} \int_{t_i}^{t_{i+1}} \psi^{\mathrm{T}}(t) \left(-\frac{\partial v}{\partial t} + f(v, q) \right) \mathrm{d}t, \tag{A.6}$$

which is the sum of the original functional and the integral over time of the differential constraint multiplied by a vector function $\psi(t)$ of Lagrangian multipliers. This expanded functional has the same extrema as the original functional (A.2).

We apply techniques of optimal control theory and variational theory [22,37], and derive analytically the necessary conditions of optimality of the 1st order, using stationary condition as described below. Such necessary conditions for the minimum of functional L, together with system (A.1) constitute a closed system of equations, the solution of which provides the optimal set of parameters.

Necessary conditions include the constraint that ψ obey the system of differential equations for Lagrangian multipliers given by

$$\frac{\partial f}{\partial v}\psi + \frac{\partial \psi}{\partial t} = 0 \qquad \forall\, t \in [t_i, t_{i+1}). \tag{A.7}$$

ψ will typically have a discontinuity at time points where data is available. At these points the system must obey the constraints

$$\frac{\partial F}{\partial v_i} - \psi(t_i - 0) + \psi(t_i + 0) = 0, \qquad i = 1, \ldots, J-1, \tag{A.8}$$

with the condition that at the end of the time interval,

$$\frac{\partial F}{\partial v_J} - \psi(t_J) = 0. \tag{A.9}$$

The stationary condition can be written now as

$$\zeta(v, q) = 0, \tag{A.10}$$

where ζ is the Lagrangian gradient given by

$$\zeta(v, q) = \int_{t_0}^{t_J} \psi^{\mathrm{T}} \frac{\partial f}{\partial q}\, \mathrm{d}t + \frac{\partial P}{\partial q}. \tag{A.11}$$

The system of differential and integral equations which specifies the necessary conditions for an extremum cannot be solved analytically. We solve it numerically using the following algorithm [21]:

1. Numerically integrate system (A.1).
2. Numerically integrate system (A.7) in reverse time from t_J to t_0 using (A.9) and (A.8).
3. Calculate the Lagrangian gradient $\zeta^k = \zeta(v^k, q^k)$ using (A.11), where v^k denotes a solution of system (A.1) obtained during the kth iteration of the algorithm using the components of vector q^k as parameters.
4. Calculate a new set of parameters q^{k+1} as follows:

$$q^{k+1} = q^k - \alpha^k \zeta^k, \tag{A.12}$$

where $\alpha^k > 0$ is a parameter chosen to minimize the value of functional (A.2) at each step.

Steps 1–4 are repeated until the value of the functional F stops decreasing. During the calculations described here, this assessment was made interactively. When F was deemed to stop decreasing, the program run was stopped, and changes were made in α, γ, or in the type of parameter transformation (A.4) or (A.5). Occasionally small random perturbations were made in the parameters themselves. When no changes of this type improved the score, the run was assumed to reach the optimum. The set of optimal parameters q^N obtained at the last iteration N of the algorithm defines the solution of the problem.

A.2. Application to the gene network model

The spatial domain contains spatially discrete gene expression data in four nuclei at cycle 11, 8 at cycle 12, 16 at cycle 13 and 32 at cycle 14A. The gap gene data used for fits is the central 32 nuclei of the data set described [17], supplemented with *eve* data [5]. This dataset contains data on *eve* expression at two time moments ($t = 29.3$ and 46.7) during cycle 13, with no gap gene data at those times. It also contains data for *eve* and the four gap genes modelled at the time points $t = 56$ and 88. At the end of cycle 14A ($t = 88$), gap gene data is missing for the anterior most four nuclei. At $t = 72$ there is data for the four modelled gap genes. The Bcd protein is modelled as a time invariant exponential [38]. We use the same data described above to fit the gene network with a mutated *eve* gene, to take into account that mutation in *eve* does not alter gap gene expression patterns [35]. The gene network with mutated *eve* is described by system (1) with $T^{a,eve} = 0$, where a enumerates the gap genes. Expression data for *hb* (at cycle 11) was used for initial conditions.

The optimization algorithm described above is valid for systems of ODEs. We discretize PDEs (1) using a grid with M points in the spatial domain [0,1] and treat the resulting discretized equations as a system of ODEs. The OSDA is then applied to these ODEs to find parameters.

The number M of grid points in the equations is constant over time, while the number of nuclei in the embryo is doubled after each successive nuclear division. This implies that we cannot associate a grid point with a nucleus and, hence, need to interpolate gene expression data depending on grid size for fitting the model. Therefore, we use linear interpolation to obtain a data set that has M spatial points at each time moment. Numerical experiments have been performed with data interpolated using splines of various orders, but results have not shown qualitative differences with respect to the linear case.

The discretized equations depend on the spatial grid size $s = 1/(M - 1)$. On one hand, s should be small enough that the numerical approximation is valid. On the other hand, a small value of s leads to a larger value of M and thus to more interpolated data points in the gene expression dataset used for fitting. Moreover, small values of s increase computational time, that becomes critical when a large number of computations needs to be performed. Here, we have chosen $M = 32$ during the optimization for the following reasons. First, there are 32 nuclei at cycle 14A and therefore, interpolation is only necessary for cleavage cycles before cycle 14A (see below). Second, we have shown numerically that the finite difference approximation on a grid with 32 spatial points is valid. The difference between numerical solutions is negligible when M tends from 32 to infinity.

The quality functional (2) will be written in the form (A.2) after the following consideration to be taken into account. The components of vectors v_i and y_i in (A.2) are the concentrations of proteins of all genes in the

Table A.1
Values of synthesis rate R^a, promoter threshold h^a, protein half lives terms $\log 2/\lambda^a$, and diffusion coefficient in the mitosis models A–C, and in the hybrid model (H)

	A	B	C	H
R^{Kr}	0.330	0.468	0.076	0.578
R^{hb}	1.225	0.932	0.159	1.125
R^{gt}	0.569	0.699	0.167	0.981
R^{kni}	1.144	2.274	0.082	1.125
R^{eve}	0.230	0.615	0.190	0.839
h^{Kr}	−1.901	−1.950	−2.367	−0.904
h^{hb}	−18.440	−18.264	−34.648	−13.592
h^{gt}	−0.905	−0.840	−0.934	−1.172
h^{kni}	4.498	4.584	8.828	2.448
h^{eve}	−3.217	−3.473	−8.758	−4.983
$\log 2/\lambda^{Kr}$	27.588	21.150	21.418	14.104
$\log 2/\lambda^{hb}$	5.721	8.770	5.531	7.031
$\log 2/\lambda^{gt}$	11.938	11.438	4.319	8.806
$\log 2/\lambda^{kni}$	10.531	4.931	13.415	8.886
$\log 2/\lambda^{eve}$	19.530	7.373	4.403	6.000
D^{Kr}	0.107	0.171	0.230	0.160
D^{hb}	0.023	0.026	0.044	0.004
D^{gt}	0.070	0.061	0.237	0.049
D^{kni}	0.009	0.015	0.047	0.013
D^{eve}	1.6×10^{-11}	2.3×10^{-9}	1.9×10^{-10}	4.2×10^{-9}

In the diffusion part of the table, each box for models A–C contains a value $(M-1)^2 D^a = 961 D^a$ (for $M = 32$) for gene a. In the hybrid model box the values D^a are presented for this model in cleavage cycle 14.

Table A.2
The interconnectivity matrix T^{ab} and maternal gene regulatory input m^a (the column under *bcd*) in the mitosis models A–C, and in the hybrid model (H)

	Kr	*hb*	*gt*	*kni*	*eve*	*bcd*	
Kr	0.437	−0.805	−0.565	−2.110	0.006	1.884	A
	0.553	−0.842	−0.508	−2.398	−0.00008	2.012	B
	0.906	−3.226	−0.974	−3.098	−0.029	2.852	C
	0.389	−1.046	−0.534	−2.169	0.005	2.054	H
hb	0.078	0.185	−1.174	−0.649	0.003	7.591	A
	−0.038	0.167	−1.201	−1.042	−0.011	7.914	B
	0.044	0.196	−1.989	−0.243	0.034	14.847	C
	−0.012	0.145	−0.692	0.063	0.002	5.734	H
gt	−2.459	−0.329	0.317	−0.359	0.003	0.983	A
	−2.333	−0.323	0.243	−0.358	−0.0008	1.018	B
	−5.126	−0.702	0.346	−0.082	−0.039	2.474	C
	−2.605	−0.434	0.162	−0.353	0.002	1.719	H
kni	−0.292	−4.129	−1.946	−0.088	0.003	−1.920	A
	−0.288	−3.355	−1.710	−0.104	0.00001	−1.938	B
	−0.678	−8.587	−3.826	−0.080	−0.013	−3.529	C
	−0.369	−4.113	−1.386	−0.070	0.003	−0.192	H
eve	−2.546	−3.462	−2.587	−1.700	0.059	15.046	A
	−2.616	−3.280	−2.573	−1.676	0.024	16.302	B
	−4.002	−6.657	−4.261	−2.483	−0.044	28.266	C
	−2.282	−3.101	−1.835	−1.373	0.037	14.914	H

Each box in the table contains four values of T^{ab} or m^a obtained in the models; the models are ordered from top (A) to bottom (H), as indicated at the rightmost column.

network for two genotypes at all grid points at time t_i calculated in the framework of the model given and obtained experimentally.

A.3. Results of optimization

We present here the parameters obtained by the optimization procedure described in the previous sections. Tables A.1 and A.2 show sets of optimal parameters for models A–C compared to optimal parameters of the hybrid model (H) described in [5].

We found that the models A–C are generally consistent with what was found in previous study of the hybrid model in [5]. Although the values of parameters can differ significantly from model to model, the signs of the regulatory parameters T^{ab} and m^a (and hence the gene network topology) are keeping constant among models. This implies that models A–C and the hybrid model all describe gene networks, which are equivalent qualitatively to each other. In these networks, we observe that gap genes are generally activated by Bcd, show repressive cross interactions and positive autoregulatory terms. For example, all models show that *Kr* is repressed by all other gap genes and is activated by Bcd as well as through self-activation (Table A.2). Moreover, *eve* is activated by Bcd and repressed by the gap genes, whereas its inputs to the gap genes are close to zero in all models. The only inconsistency between models A–C and the hybrid model is in regulation of *hb* by Kni, which is small positive in the hybrid model, whereas it is negative in models A–C.

References

[1] A.M. Turing, Trans. R. Soc. London B 237 (1952) 37–72.
[2] R. Thomas, J. Theor. Biol. 153 (1991) 1–23.
[3] R. Thom, Structural Stability and Morphogenesis, W.A. Benjamin, Reading, MA, 1975.
[4] E. Mjolsness, D.H. Sharp, J. Reinitz, J. Theor. Biol. 152 (1991) 429–453.
[5] J. Reinitz, D.H. Sharp, Mech. Dev. 49 (1995) 133–158.
[6] C. Nusslein-Volhard, E. Wieschaus, Nature 287 (1980) 795–801.
[7] C. Nusslein-Volhard, E. Wieschaus, H. Kluding, Roux Arch. Dev. Biol. 193 (1984) 267–282.
[8] G. Jurgens, E. Wieschaus, C. Nusslein-Volhard, H. Kluding, Roux Arch. Dev. Biol. 193 (1984) 283–295.
[9] E. Wieschaus, C. Nusslein-Volhard, G. Jurgens, Roux Arch. Dev. Biol. 193 (1984) 296–307.
[10] C. Nusslein-Volhard, H.G. Frohnhofer, R. Lehmann, Science 238 (1987) 1675–1687.
[11] M. Akam, Development 101 (1987) 1–22.
[12] P.W. Ingham, Nature 335 (1988) 25–34.
[13] P.A. Lawrence, The Making of a Fly, Blackwell Sci. Publ., Oxford, UK, 1992.
[14] P.T. Merrill, D. Sweeton, E. Wieschaus, Development 104 (1988) 495–509.
[15] E. Wieschaus, D. Sweeton, Development 104 (1988) 483–493.
[16] S.B. Carroll, M.P. Scott, Cell 45 (1986) 113–126.
[17] J. Reinitz, E. Mjolsness, D.H. Sharp, J. Exp. Zool. 271 (1995) 47–56.
[18] J. Reinitz, D. Kosman, C.E. Vanario-Alonso, D.H. Sharp, Dev. Gen. 23 (1998) 11–27.
[19] V.V. Gursky, J. Reinitz, A.M. Samsonov, Chaos 17 (2001) 3–12.
[20] V.A. Foe, B.M. Alberts, J. Cell. Sci. 61 (1983) 31–70.
[21] K.N. Kozlov, A.M. Samsonov, Tech. Phys. 48 (2003) 1364–1371.
[22] A.E. Bryson, Y.C. Ho, Applied Optimal Control, Taylor and Francis, 1975.
[23] L. Sánchez, D. Thieffry, J. Theor. Biol. 211 (2001) 115–141.
[24] G.K. Yasuda, J. Baker, G. Schubiger, Genes Dev. 5 (1991) 1800–1812.
[25] T.C. Lacalli, D.A. Wilkinson, L.G. Harrison, Development 104 (1988) 105–113.
[26] H. Meinhardt, Models of Biological Pattern Formation, Academic Press, New York, 1982.
[27] A. Hunding, S.A. Kauffman, B.C. Goodwin, J. Theor. Biol. 145 (1990) 369–384.
[28] T.C. Lacalli, J. Theor. Biol. 144 (1990) 171–194.
[29] H. Meinhardt, J. Cell. Sci. Supp. 4 (1986) 357–381.
[30] H. Meinhardt, Development 104 (1988) 95–110.

[31] B.N. Nagorcka, J. Theor. Biol. 132 (1988) 277–306.
[32] H. Meinhardt, J. Embryol. Exp. Morphol. 76 (1983) 115–137.
[33] B.A. Parr, A.P. McMahon, Nature 374 (1995) 350–353.
[34] R. Warrior, M. Levine, Development 110 (1990) 759–767.
[35] M. Frasch, M. Levine, Genes Dev. 1 (1987) 981–995.
[36] D. Kosman, S. Small, J. Reinitz, Dev. Genes E 208 (1998) 290–294.
[37] K.N. Lurie, Applied Optimal Control Theory of Distributed Systems, Plenum, New York, 1993.
[38] W. Driever, C. Nusslein-Volhard, A gradient of bicoid protein in *Drosophila* embryos. Cell 54 (1988) 83–93.